DOCUMENT RESUME

ED 184 052 CS 005 222

AUTHOR Vaughan, Joseph L., Jr., Ed.; Gaus, Paula J., Ed.
TITLE Research on Reading in Secondary Schools: A Semi-Annual Report, Volume 2, Number 2. Monograph Number Three.
INSTITUTION Arizona Univ., Tucson.
PUB DATE 79
NOTE 77p.; Best Copy Available

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS *Content Area Reading; Educational Research; Graphs; *Learning Theories; Literature Reviews; Readability Formulas; *Reading Comprehension; *Reading Research; *Reading Skills; Secondary Education; *Study Guides; Teaching Methods
IDENTIFIERS Ausubel (David P); *Dale Chall Readability Formula

ABSTRACT

This monograph is the third in a series of semiannual reports of research related to reading in secondary schools. The first section of the monograph is made up of three research studies on the validation of teaching procedures to facilitate students' comprehension of content area texts, the graph-reading abilities of seventh grade students, and a comparison of three editions and the sample frequency of the Dale-Chall Readability Formula. The second section contains literature reviews of study guides as aids to learning and of research related to reflective reaction activities, based on David Ausubel's theory of learning. (AEA)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED184052

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

RESEARCH ON READING IN SECONDARY SCHOOLS

A Semi-Annual Report

Volume 2, Number 2

Edited by

Joseph L. Vaughan, Jr.
Paula J. Gaus

Spring, 1979

Reading Dept.
University of Arizona

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Joseph Vaughan

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CS005222

BEST AVAILABLE COPY

EDITORIAL ADVISORY BOARD

Editors:

Joseph L. Vaughan, Jr.
Paula J. Gaus

Board:

Wilbur S. Ames
Patricia L. Anders
Jill LaZansky
Jim Rosenthal

.....

A special thanks is due to Kenneth J. Smith, Head of the Reading Department for his support in producing this monograph. Further, we would like to thank our support staff, especially Lil Tolley, for all their assistance in preparing and compiling this issue.

.....

All materials in this monograph have been copyrighted © 1979. All materials reserved. No part of any of this publication may be printed, reproduced, or transmitted in any form or by any means without prior permission from the editors.

Table of Contents

Page

Editorial Advisory Board i

Table of Contents ii

Purchase of Monographs iii

Submission of Manuscripts iv

Section One: Research Studies

Validating Teaching Procedures Designed to Facilitate Students' Comprehension of Content Area Texts
Michael F. Graves, Associate Professor
University of Minnesota 1

Graph-Reading Abilities of Seventh Grade Students
Beverly V. Roller, Jefferson County Schools
Jefferson County, Colorado 16

A Comparison of Three Editions and Sample Frequency of the Dale-Chall Readability Formula
Robert T. Williams, Ed.D., Associate Professor
Department of Education
Colorado State University 39

Section Two: Reviews of Literature

Study Guides as Aids to Learning: A Review of Research
Jodi Grant, Ph.D. Candidate
Reading Dept, University of Arizona 51

Research Related to Reflective Reaction Activities
Barbara Flores, Ph.D. Candidate
Reading Dept, University of Arizona 62

Future Issues 71

PURCHASE OF MONOGRAPH

This monograph is the third in a series of semi-annual reports of research related to reading in secondary schools. This monograph has been funded by the Reading Department of the University of Arizona for the primary purpose of disseminating research findings to professionals in reading who are interested in improving the reading ability of adolescents.

The first issue of this monograph series was published in April, 1978, and may be obtained for $2.50. Subsequent issues, including the current issue, are available at a cost of $3.50. Checks should be made payable to Reading Research Monograph and forwarded to the University of Arizona, Tucson, Arizona 85721.

The requests for copies of these monographs have been overwhelming and we have tried to keep pace with requests. Unfortunately, we have not done this as well as we'd like, but we are trying our best. We have a limited number of the first two issues remaining and we are trying to expedite our dissemination process. If you want back issues and/or want to be added to our mailing list, please let us know.

SUBMISSION OF MANUSCRIPTS

This monograph series is planned and published as a non-profit enterprise by faculty and doctoral students in the Reading Department at the University of Arizona. Each article submitted is independently reviewed by at least three members of the editorial board. Subsequent issues are planned that will include specific research studies, theoretical papers that emphasize a need for research, and additional reviews of literature. Articles that deal with any of these areas are being actively solicited if they are (1) of interest to reading professionals on the secondary level, (2) are primarily directed to research, and (3) have not previously been published elsewhere.

One unique aspect of this series is that the editorial board is particularly interested in receiving papers from doctoral students and from assistant professors. It is our belief that many of the professionals in these categories do not publish their research findings because of a perceived and existing difficulty "breaking into the publishing game." We hope to encourage "first-time" authors by providing such an opportunity and encourage such prospective authors to submit papers or questions to the editors.

The response to our effort has been most gratifying. As you may well appreciate, the production and dissemination of a new publication is hardly an easy task. In fact, as we succeed, the task becomes increasingly more complex. Our future success depends upon the degree to which our colleagues join us in generating research and in reporting their findings. It is to this continuing end that this third monograph is dedicated.

J.L.V., Jr.
P.J.G.

Validating Teaching Procedures Designed to Facilitate Secondary Students' Comprehension of Content Area Texts *

Michael F. Graves
sociate Professor
University of Minnesota

For the past several years, a research group at the University of Minnesota has been conducting a series of classroom experiments which attempt to validate specific teaching procedures useful to secondary content teachers in facilitating their students' reading of class materials. At the present time, six studies have been completed (Bender, 1975; Clark, 1978; Miller, 1978; Mrocek, 1978; Palmer, 1978; Pearson, 1978); two have been submitted for publication (Graves & Bender, 1979; Graves & Clark, 1979); several are in progress (Graves, Patberg, Henning, & Palmer; Graves & Palmer; Hood); and two general reports on the line of research have been presented (Graves, 1977, 1978a). This paper is a revised version of the 1978 general report and attempts to give a rather complete description of the line of research. More specifically, this paper briefly describes the general characteristics of the research and presents the results of several completed studies and some in progress, and comments on what has been done and our future plans.

---

*The present paper is a substantially revised version of a paper presented at the 1978 NCTE Preconvention Conference on Research held in Kansas City, Missouri.

General Characteristics of the Studies

The studies can be described as having seven general characteristics. First, the teaching procedures to be examined are ones that can appropriately and feasibly be used by regular content teachers with five classes of 30 or so students. This restriction is obviously an arbitrary choice, and many other sorts of classroom research also need to be done. However, I am satisfied with the restriction and believe the sort of work it leads to is important. At the same time, two qualifications are needed here. One is that saying that the procedures are feasible for regular classroom teachers does not mean that they require no work beyond what the teacher is currently doing. Unfortunately, I know of few procedures to facilitate students' reading that do not require work on the part of the teacher. The other qualification is that some of the procedures investigated will be of use to only some, perhaps few, classroom teachers. For example, in one recent study (Clark, 1978), we looked at the effect of simplifying vocabulary and syntax on students' comprehension. Such modification of texts is, of course, extremely time-consuming, but the procedures can be used by at least some teachers some of the time.

The second general characteristic of the studies is that they are of teaching procedures which have as their principal goal the teaching of content as opposed to the teaching of reading skills. As with the other restriction, this one is obviously an arbitrary choice, and other sorts of classroom research need to be done. Again, however, I am satisfied with this choice. Also, I might add that while I previously justified this choice in terms of authority (Herber's, 1970, position that the first responsibility of a content teacher is to teach content), I can also justify it in more pragmatic terms. In seven years of teaching secondary reading courses and conducting inservice programs in the field, I

have found that almost all secondary content teachers will undertake extra work for the sake of teaching their content much more readily than they will undertake extra work to facilitate the development of reading skills.

The third general characteristic of the studies is that they are of procedures connected with specific reading selections; that is, the studies involve procedures designed to facilitate comprehension of something like Faulkner's "Two Soldiers" rather than procedures designed to generally facilitate comprehension of short stories. This is a necessary corollary of the decision to investigate procedures which teach content rather than reading skills.

The fourth characteristic of the studies is that they are of procedures appropriate for short selections - a chapter, an essay, or a short story - rather than for book length works such as novels. Although all of the work which we have done so far has been with short selections, although such work is certainly less complex than that with longer selections, and although some of the teaching procedures we investigate may be more appropriate with shorter selections than with longer ones, the restriction seems unnecessary as a defining characteristic of the work. It might, for example, be particularly interesting to investigate the successive use of several procedures while teaching a novel. At the same time, my present plans are to stay with shorter selections and hence shorter studies.

The fifth general characteristic of the studies is a corollary of the third. As a result of investigating procedures connected with specific reading selections, the gains considered are short term and specific rather than long term and general. I do not believe that this means that the research is of little value. For one thing, while the result of a teacher's using a particular teaching technique (say, for example, previewing) is that students

learn about a specific selection, the results of disseminating research showing that previewing is effective is conceivably that many teachers will use previewing to facilitate students' dealing with many selections. For another thing, such a restriction is nothing new. The research on techniques for facilitating the learning of specific selections [for example, the work of Ausubel (1960, 1968) with advance organizers; Meyer (1975, 1977) with prose analysis; and Rothkoph (1965, 1966) with adjunct questioning] have a respected place in educational research. Finally, validating teaching procedures that lead to short-term gains seems a reasonable step toward validating procedures that lead to long-term gains, a goal that educational research has frequently attempted without success.

The last two general characteristics of the studies are closely related. The first is that the work is deliberately designed to be replicable, that the process of the research is described in sufficient detail that the studies can be run again in precisely the same fashion as the first time. I believe that this feature of the work is extremely important. Attempts to replicate educational research have shown it to be notoriously unreliable. While one may not want to go as far as Carver (1978) does in his recent paper on the perils of statistical significance testing, I'm becoming increasingly skeptical of inferential statistics in a world where results significant at $p < .01$ fail to replicate.

The characteristic related to replicability is that of robustness. The studies are of low risk procedures, procedures which previous research, theory, authority, and common sense all say ought to work. Like the notion of replicability, that of doing low risk research is one that I would like to endorse very strongly. Teaching procedures which research, theory, authority, and common sense strongly suggest will work don't always work. Our work has revealed several instances in which treatments are ineffective.

Even low risk research, research of techniques which are generally agreed upon as valuable and are widely used, fail to produce consistent results. Investigation of the validity of these sorts of activities seems a more immediate concern than investigation of those that seem to be less powerful and are less widely used.

In addition to having these characteristics, the studies generally follow the guidelines set forth in Structuring Reading Activities for English Classes (Graves, Palmer, & Furniss, 1976), and they are generally consistent with what Rosenshine (1978) has termed "direct instruction."

Described below are three studies on adjunct questioning and two on previewing. The descriptions are deliberately brief, and while the designs actually included several variables, only those of major interest here are noted.

Adjunct Questioning Studies

The adjunct questioning studies are those of Clark (1978), Palmer (1977), and Pearson (1978). Subjects for Clark's study were 180 low ability tenth, eleventh, and twelfth graders. Subjects for Palmer's study were 72 low, middle, and high ability eighth graders. And subjects for Pearson's study were 89 middle and high ability eighth graders.

Materials for the Clark study were a fifth grade version and a tenth grade verson of a 2000 word article on the history of the phonograph and ten, text-dependent, multiple-choice questions with four alternatives. Materials for the Palmer and Pearson studies (Pearson's study was largely a replication of Palmer's) were a 3500 word selection on modern architecture and ten text-dependent, multiple-choice questions with four alternatives. In all studies, the selections were typed double-spaced on letter size paper, both with and without the questions inserted. In the Clark study, questions were inserted at 200 word intervals following the paragraphs in which they were answered. In the Palmer and Pearson

studies, the questions were inserted at 400 word intervals following the paragraphs in which they were answered. The questions were also typed separately for each study.

As considered here, each of the studies can be considered as having a single variable question placement with three levels - interspersed questions, final questions, no questions. The results given are averaged across all other variables. The studies were analyzed using the analysis of variance and Newman-Keuls procedures.

Subjects in each study were tested in groups. Those in the Clark study received oral instructions and were specifically told that they would be tested on the material the following day. Those in the Palmer and Pearson studies received written instructions and were told to read the selection as if they were going to be tested on it. In all studies, subjects receiving interspersed questions were told that they would be receiving questions in the text and should answer them, and subjects receiving final questions were told that they would be receiving questions just after reading the text and should answer them. All subjects received the same ten questions that were given to the interspersed and final question groups following the reading and again one week later. The results of the three studies are shown in Table One below.

TABLE ONE

Means for Three Levels of Question Placement on Clark, Palmer, and Pearson Studies

Study	Question Placement		
	Interspersed Questions	Final Questions	No Questions
Clark	7.3	6.3	5.2
Pearson	8.8	8.0	7.6
Palmer	7.4	8.1	5.6
$\bar{x}$	7.8	7.5	6.1

As can be seen from the table, in the Clark and Pearson studies, interspersed questions produced greater recall than final or no questions; final questions produced greater recall than no questions. In the Palmer study, final questions produced greater recall than interspersed or no questions; interspersed questions produced greater recall than no questions.

The analysis of variance for the Clark study indicated a significance difference ($p < .01$) due to question placement. The Newman-Keuls procedure showed that in this study interspersed questions produced significantly greater recall than final questions ($p < .05$) or no questions ($p < .01$) and that final questions produced significantly greater results than no questions ($p < .01$). The analysis of variance for the Pearson study indicated a significant difference ($p < .05$) due to question placement. The Newman-Keuls procedure showed that in this study interspersed questions produced significantly greater recall than no questions ($p < .05$) and that this was the only significant difference. The Palmer study was analyzed somewhat differently and a direct comparison of the analysis of this study with the other two is not possible; however, an analysis of variance did indicate that the difference between the recall produced by interspersed questions was not significantly different ($p = .17$) than that produced by final questions. Further, inspection of cell means suggests that the difference between the recall produced by interspersed and final questions and that produced by no questions is significant.

In summary, these three studies indicate that if we want students to recall the answers to specific questions, placing those questions within the text or immediately following the text is likely to facilitate recall. However, the studies do not indicate whether (1) interspersed questions are superior to final ones, (2) final questions are superior to interspersed ones, or (3) the effect of both sorts of questions is about equal. I will comment on the ambiguity of those results in the concluding section of this paper.

Previewing Studies

The previewing studies described here are those of Hood (in progress) and Graves and Palmer (in progress). Subjects for Hood's study were 41 tenth graders of varying abilities. Subjects for the Graves and Palmer study were 322 seventh, eighth, and ninth graders of middle and high ability.

Materials for the Hood study included two short stories averaging about 1,700 words in length and a 400 word preview for each of the stories. To assess recall, eighteen multiple-choice questions were used, six of which were factual questions answered in the preview, and six of which were inferential questions. For both the Hood study and the Graves and Palmer study, a preview was defined as an overview of a story which is written at the same level of generality as the selection and which attempts to give students specific information about plot, characters, and setting. Materials for the Graves and Palmer study included two short stories averaging 2,000 words in length, a 400 word preview for each of the stories, and eighteen multiple-choice questions of the same three types as those in the Hood study.

As considered here, each of the studies can be considered as having a single variable treatment with two levels -- preview and no preview. The results given are averaged across other variables. The results of these studies have not yet been statistically analyzed.

Subjects in both studies were tested in groups. The preview groups had the preview read aloud to them while they followed along on a typed copy, read a story, and took the test without looking back at the story. The no-preview groups simply read a story and took the test.

The results of both studies are shown in Table Two on the next page.

TABLE TWO

Means for Two Levels of Previewing
on Hood and Graves and Palmer Studies

Study	Treatment	
	Preview	No Preview
Hood	14.8	10.1
Graves and Palmer	15.0	14.5
$\bar{x}$	14.9	12.3

As can be seen from the table, in the Hood study having a preview greatly facilitated students' performance, with those students receiving the preview producing 47% more correct responses than those who did not have the preview. On the other hand, in the Graves and Palmer study having a preview did very little to facilitate students' performance, with those students receiving the preview producing only 3% more correct responses than those who did not have the preview. Moreover, in all likelihood the difference in the Hood study is statistically significant while that in the Graves and Palmer study is not.

In summary, these two studies indicate that previews sometimes work and sometimes do not work. In the section that follows I will discuss this inconsistency and comment on this line of research more generally, some things we have learned from it, and our future plans.

Discussion

My remarks here fall into three areas. The first concerns the instability of the results and the consequent importance of replication and establishing boundary conditions. The second concerns the importance of low risk research and testing common sense. And the third concerns methodological issues.

Importance of Replication

My motivation for stressing the importance of replication is obvious. Our attempts at "loose replication" (doing generally similar experiments) have not been successful. What appears to produce a huge effect in one study (the presence or absence of a preview, for example) may have no effect in another. I think that we need to move from loose replication to true replication if we are to achieve stability of results. At the same time, it must be recognized that doing true replications presents obvious practical problems. Journals seldom publish replications, and replications are not necessarily appropriate or sufficient for purposes of completing research done as part of a graduate degree. (My greatest source of this sort of work is MA papers). I know of no solution to the latter problem; however, the former problem might be somewhat alleviated if journals would make replication a requirement for any studies where replication is feasible. (Ed. Note: This monograph series welcomes replication.)

Importance of Boundary Conditions

My motivation for stressing the importance of boundary conditions is equally obvious. Our procedures seem to work under some conditions but not under others. While it is my intention to work first with the most robust procedures, it appears that few, if any, procedures are as robust as we would like or as many people consider

them to be. Thus, we need to discover under just what conditions (age of students, ability of students, characteristics of material, not to mention characteristics of teacher) various procedures are effective. It may be, for example, that the discrepancies in the results of the questioning studies can be at least partially explained by taking into account the ability of the students tested. The Clark study included only low ability students; the Palmer study included both low and high ability; and the Pearson study examined only middle and high ability. The Clark study showed the greatest effect of questioning, the Palmer study somewhat less effect, and the Pearson study much less effect.

My interest in doing low risk research is apparent from the comments about robustness made above, and I add this as an additional point only for emphasis. Procedures which common sense, available theory, and authority suggest will work frequently don't work. The first job, either in terms of likelihood of payoff or in terms of getting rid of frequently used ineffective techniques, is to test these procedures.

Methodological Concerns

My major methodological concerns can be briefly noted. I am concerned with our dependent measures, multiple-choice tests. There is the obvious possibility of using some sort of protocol analysis which would be more in keeping with much of the present work in both education and psychology. At the same time, such dependent measures come of an enormous cost, and I am not sure what they would gain. I am also concerned with the fact that we haven't asked teachers and students to introspect about the procedures. Asking such questions costs little, and I intend to begin talking to teachers and students more. Finally, there is the matter of the complexity of the designs. As I noted, the descriptions of the designs presented above were greatly simplified. For example,

the Graves and Palmer previewing study in progress employs a 3 x 3 x 2 x 2 x 3 design with the variables being grade (7th, 8th, 9th), ability (high, middle, low), treatment (preview, no preview), story/difficulty (6th grade, 9th grade), and question type (factual in preview, factual not in preview, inferential not in preview). While this certainly isn't complex when judged by the standards of university researchers, getting it set up in a school such that it doesn't terribly disrupt the school and isn't hopelessly confounded is no mean task. In fact, such designs preclude wide scale replication at multiple sites. In light of this, I have considered the desirability of moving to simple one factor, two level designs (see Graves, 1978b for examples of these).

By way of conclusion I will simply say that I distinctly want criticism of our work and would welcome comments. We have the materials available so that at least some of these studies can be replicated. I would also welcome correspondence from those interested in replicating one or more of the studies.

(Ed. Note: Dr. Graves can be contacted at the College of Education, University of Minnesota, Minneapolis, Minn. 55455)

References

Ausubel, D. P. Educational psychology: a cognitive view. New York: Holt, Rinehart & Winston, 1968.

Ausubel, D. P. The use of advance organizers in learning and retention of meaningful verbal material. Journal of Educational Psychology, 1960, 51, 267-272.

Bender, S. D. Teaching vocabulary in high school: a classroom experiment. Unpublished master's paper. University of Minnesota, 1975.

Carver, R. P. The case against statistical significance testing. Harvard Educational Review, 1978, 48, 378-398.

Clark, D. Effects of adjunct questions on high school low achievers. Unpublished master's paper. University of Minnesota, 1978.

Graves, M. F. Validating specific teaching procedures to be used by secondary content area teachers. Paper presented at the NCTE Preconvention Conference on Research. New York, November, 1977.

Graves, M. F. Validating teaching procedures designed to facilitate secondary students' reading of content area materials. Paper presented at NCTE Preconvention Conference on Research. Kansas City, November, 1978a.

Graves, M. F. Classroom research as a vehicle for inservice education. Report prepared for the University of Minnesota/St. Paul Public Schools Cycle XII Teacher Corps, October, 1978b.

Graves, M. F. and Bender S. Preteaching vocabulary to secondary students: a classroom experiment. Submitted to Minnesota English Journal, January, 1979.

Graves, M. F. and Clark, D. L. The effect of adjunct questions on high school low-achievers' reading comprehension. Submitted to the Journal of Reading, January, 1979.

Graves, M. F. and Palmer, R. J. The effect of previewing on the recall of junior high school students. University of Minnesota, in progress.

Graves, M. F., Palmer, R. J. and Furniss, D. W. Structuring reading activities for English classes. Champaign, Illinois: NCTE, 1976.

Graves, M. F., Patberg, J. P., Henning, M., and Palmer, R. J. An investigation of three methods of preteaching vocabulary. University of Minnesota and University of Toledo, in progress.

Herber, H. L. Teaching reading in content areas. New Jersey: Prentice-Hall, 1970.

Hood, M. The effect of previewing on the recall of high school students. Unpublished master's paper. University of Minnesota, in progress.

Meyer, B. J. F. What is remembered in prose: a function of passage structure. In R. O. Freedle (Ed.), Discourse production and comprehension. Norwood, N.J.: Ablex Publishing Company, 1977.

Meyer, B. J. F. The organization of prose and its effects on memory. Amsterdam: North-Holland Publishing Co., 1975.

Miller, L. A comparison of three methods of improving reading comprehension in junior high schools: Vocabulary simplification, preteaching vocabulary, and prereading discussion. Unpublished master's paper. University of Minnesota, 1978.

Mrocek, D. The effectiveness of two games to teach vocabulary to secondary students. Unpublished master's paper. University of Minnesota, 1978.

Palmer, R. J. The effect of postquestion type and postquestion placement on the prose learning of high and low ability seventh graders. Unbpublished master's paper. University of Minnesota, 1977.

Pearson, B. The effect of postquestion type, postquestion placement, and immediate and delayed retention posttests on the prose learning of high and low ability 8th graders. Unpublished master's paper. University of Minnesota, 1978.

Rosenshine, B. Instructional principles in direct instruction. Paper presented at the National Reading Conference, St. Petersburg, Florida, December, 1978.

Rothkoph, E. Z. Learning from written instructive materials: an exploration of the control of inspection behavior by testlike events. American Educational Research Journal, 1966, 3, 241-249.

Rothkoph, E. Z. Some theoretical and experimental approaches to problems in written instruction. In J. D. Krumboltz (Ed.), Learning and the educational process. Chicago: Rand-McNally, 1965.

Graph-Reading Abilities of Seventh Grade Students

Beverly V. Roller
Jefferson County Schools
Jefferson County, Colorado

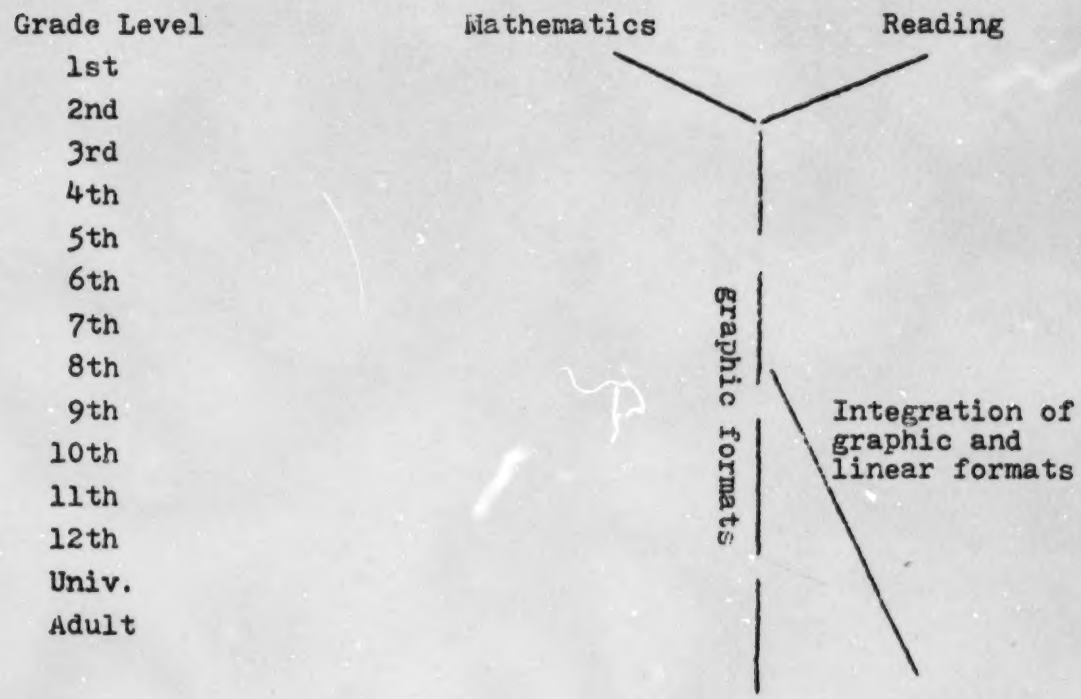
Graphs are one of the multidisciplinary devices of the secondary curriculum, a visual method of presenting large amounts of compacted information, illustrating relationships and comparisons. Graph reading is usually classed as a study skill; responsibility for initial instruction is often assigned to mathematics and reading. Science and social studies are seen as the major content area which require application of graph reading skills. By the ninth grade, consumer mathematics, (a sub-content area of survival or functional reading) is also a heavy user of graphs.

The author who decides to use a graph as an illustration often does so because a great deal of numerical data can be pictured in a compact space, sometimes saving the author the writing of lengthy paragraphs of details. Because so much information is presented in a condensed format, graphs, by their nature, represent a difficult reading task.

Summary of Graph Reading Literature

The literature of graph reading is in bits and pieces and appears to parallel the changing requirements and expectations placed upon readers at various grade levels in school and into adulthood. Figure One depicts the visualization developed from a number of references which cite and/or discuss graphs.

Figure One: Visualization of Graph Reading Research



During the early primary years, the issues center around the age at which instruction should begin (Strickland, 1938). Line, bar, circle and picture graphs are distinctive and the formats are easily recognizable. However, there is a question about the age and/or level at which children achieve the concept of relationship which underlies all graphs. This, then raises the issues (1) are graphs useable to teach the concept of relationship, and (2) can children be said to "read" graphs if they do not understand the concept of relationship?

During the intermediate years, the issue is graph instruction. Without exposure and instruction, learners do not understand the graphic formats. How each individual responds to graph instruction appears to be directly related to, and limited by, the student's understanding and mastery of the mathematical sequence. By the time the majority of students reach the eighth grade, they appear to be able to read, understand, and make interpretations which are a synthesis of reading, mathematics, and graph instruction (Buswell, 1937; Otto, Kamm, Weible, 1972).

At about sixth grade, there is a shift in the expectations placed upon readers. Readers are expected to use and integrate information presented in a narrative or expository text with information presented in the graphic format. This expectation parallels an increasing emphasis upon content area and textbook reading. Graphs encountered in textbooks often require the reader to stop, study the information, and then, return to the text. This stopping and starting requires a flexibility on the part of the reader and several studies have noted the reluctance of readers to interrupt the flow of reading a linear text.

Vernon (1946, 1950, 1951, 1953, 1954) did the most in-depth studies with learners ranging from early adolescence through adulthood and noticed that younger learners were content with isolated facts while older learners were more likely to remember

general arguments. Her work suggests that there is a developmental sequence surrounding the integration of textual and graphic information. Turner (1974), McCallister (1930), and Shores (1943) studied the sixth to ninth grade range, and there are hints in their reports that readers are dependent upon teacher intervention and direction to make sense out of the combined presentations and, even, to focus attention upon information in the graphic format. Vernon (1950) estimated that the reader must be at the level of a university student before being independent of teacher intervention and capable of integrating text and graph as well as other types of illustrations.

Research has tended to center on whether a graphic aid contributes to the comprehension of concepts carried primarily by the text. However, the study of Eggen, Kauchak and Kirk (1978) with university students raises the question: to what extent do two types of information presentation interfere with the comprehension of the reader? Their findings, combined with other sources, suggest that when graphs are the focus, the text will interfere; and when text is the focus, graphs will interfere. It appears that researchers, as well as authors, must be clear as to the type of presentation of information which is expected to carry the logic and weight of the major point the reader is to comprehend. There are, then two research questions:

1. If the major point to be made with the reader is carried by the text, to what extent does a graphic aid contribute to the reader's comprehension?

2. If the major point to be made with the reader is contained in a graphic format, to what extent does narrative text contribute to, or interfere with, the reader's comprehension?

There is one major problem when studying graphs, whether it is by students in a classroom or researchers studying students as they read graphs. The problem is that every graph is unique. Graphs in materials which readers may encounter in the course of daily reading activities show a wide variation in the illustrator's art, particularly at the secondary and adult levels.

For example, line graphs depend upon a figure and ground relationship between the information-bearing line(s) and a grid. In terms of grid lines, it is possible to draw a line graph in at least four ways: (1) complete grid, both horizontal and vertical lines; (2) horizontal grid lines only; (3) vertical grid lines only; and (4) no grid lines. It is further possible to affect that figure-ground relationship by making the grid lines finer or of the same intensity as the information-bearing lines. The whole area of graph readability and understandability is clouded because it is impossible to generalize from one line graph to all line graphs.

Graph difficulty appears to depend upon three major factors: primarily (1) the number of comparisons pictured, and then (2) the titling and labeling of those comparisons, and (3) the measurement scales used to display those comparisons. As a rule of thumb, a one-line graph is probably easier than a two-line graph. However, if we put two two-line graphs side by side, are they of equal difficulty? The answer may depend on whether or not the information lines interact, how those lines are labeled, etc. There is infinite variation possible within the graphic format, and this variation contributes to making graphs as interesting and "slippery" topic for investigation.

There is a subtle theme of reader avoidance running through the literature of graph reading. Vernon (1954) saw the problem

for the early adolescent in terms of the discrepancies between the subject matter pictured and the experential background of the reader. From one point of view it could be suggested that graphs which do not match the experience background of the reader are a waste of time and should not be used. If the viewpoint is taken that experiences can be extended by reading graphs, then there appears to be reason to continue the instructional process. Since there seems little doubt that authors, editors and publishers will continue to include graphs as illustrations, the teacher is faced with decisions as to the use of such illustrations, based on an intimate knowledge of the students in his or her classroom.

The seventh grade appears to be the level at which there is the most definite "flip" in the curriculum and in the teacher's attitude toward reading. Secondary content teachers, whether they verbalize it or not, expect reading activities to extend the experiential background of the reader. Intellectually, there is no quarrel with this position.

On the feeling level, however, there are a number of questions. How do the students feel about being asked to respond to graphs which picture information on topics removed from their everyday lives? Do they like them? Do they feel successful? Do they feel pressured? In looking for answers to these questions it was noted that there are almost no directly reported interview data. Hence there appeared to be a place for a study with seventh graders, questioning how they view the graph reading task.

Purposes of the Study

One purpose of the study was to determine how well seventh grade students can apply the skills of graph reading, graph interpretation, and graph summarization in social studies, science

and consumer mathematics situations, with and without a supporting text. A second purpose was to study the relationship of achievement in reading comprehension and mathematics computation to graph skills. A third purpose was to determine the feeling and perceptions of students when faced with the graphic format.

Procedure

Two testing instruments, using a multiple-choice item format, were developed and informally named Text Graphs and Isolated Graphs. The reliability coefficients were .60 and .72 respectively, determined by Cronbach's Alpha, a test of internal consistency. Text Graphs consisted of two textual selections from social studies and science textbooks. The selections contained approximately 2420 words (Total), two sets of two circle graphs and one simple line graph, accompanied by fifteen comprehension items. The test labeled Isolated Graphs consisted of eight graphs, the three from textbooks (Test Graphs) plus five additional simple and complex bar and line graphs taken from the business section of a popular periodical. These eight graphs were accompanied by thirty-eight comprehension items. Three graphs and ten items on the two instruments were identical and were used to compare the effects of graph presentation within a text and in isolation.

The multiple-choice items, called graph items, required subjects to use skills defined as reading, interpretation and summarization:

1. Graph reading involved use of directly stated information, including estimating amounts between grid lines on bar and line graphs and relative size on circle graphs.

2. Graph interpretation involved use of computation (limited to addition and/or subtraction) to achieve interpolated information, or two comparisons of areas as on circle graphs.
3. Graph summarization involved recognition of the statement which best described the total graphic presentation.

In addition, some items were classified as non-defined and were not used in the statistical analysis.

Two hundred sixty-eight seventh graders, 136 boys and 132 girls, in three junior high schools, representing regional areas of a large 80,000 student school system were tested, 119 taking Text Graphs and 149 taking Isolated Graphs. Four groups of fifteen subjects each were identified as random group administration, random individual administration, high scorers, and low scorers on Isolated Graphs. These subjects were interviewed as to how their answers had been obtained. Interview subjects were also asked to identify graphs which they perceived to be easy or difficult, and to indicate whether they liked graphs, where they had learned about graphs, and whether they ignored or paid attention to graphs when reading in a textbook.

Following identification of subjects in the study, reading comprehension and mathematics computation scores were located. The subtests of the Comprehensive Test of Basic Skills had been administered to most of the subjects as middle-of-the-year sixth graders.

Results

To compare the effects of graph presentation with or without a supporting text (the independent variables), the reading comprehension and mathematics computation scores of all subjects were entered into the Analysis of Covariance as covariates.

Responses to four reading, three interpretation and three summarization items on the two instruments (Text Graphs and Isolated Graphs) were the dependent variables. The Analysis of Covariance revealed significant differences favoring isolated presentation of graphs for items identified as reading and interpretation, but no significant differences were seen for summarization items.

SUMMARY OF ANALYSIS OF COVARIANCE COMPARING TEXTUAL AND ISOLATED GRAPH PRESENTATION VARIATION BY READING, INTERPRETATION AND SUMMARIZATION ITEMS WITH READING COMPREHENSION AND MATHEMATICS COMPUTATION SCORES (CTBS) AS COVARIATES

	Textual Graph Presentation N=119		Isolated Graph Presentation N=149		
Item Categories	Mean	Deviation	Mean	Deviation	Significance
Reading	.565	.255	.693	.222	.001
Interpretation	.385	.227	.479	.208	.001
Summarization	.592	.294	.629	.313	.356

This finding is supportive of other researchers who have found that text and graphic information are not commonly merged in the mind of the reader. Vernon (1951) established that a text with graphs is more difficult than a test without graphs, an effect noted for graph reading and graph interpretation questions. The lack of differences for summarization items appears to repeat a similar finding of Eggen, Kauchak and Kirk (1978).

Correlations

The grade equivalent means of all subjects in the study were 8.8 for reading comprehension and 7.8 for mathematics computation. To determine whether the reading comprehension and mathematics computation scores would predict how subject would respond to graph items on both instruments, a Pearson Product Moment Correlation was computed. Although it would appear that reading comprehension and mathematics computation scores (CTBS) are related to graph items of reading, interpretation and summarization, the relationships are weak.

CORRELATION OF COVARIATES AND ITEM CATEGORIES FOR 268 SUBJECTS

Variable Pairs	Coefficient	Significance
Graph Reading - Reading Comprehension	.3048	.001
Graph Reading - Mathematics Computation	.1433	.009
Graph Interpretation - Reading Comprehension	.2187	.001
Graph Interpretation - Mathematics Computation	.1418	.010
Graph Summarization - Reading Comprehension	.3027	.001
Graph Summarization - Mathematics Computation	.1229	.020

This finding suggests the inadvisability of grouping students for instructional or research purposes related to graphs, based on standardized achievement testing, (and may explain the somewhat erratic findings of Turner, 1974). Although the correlation revealed statistical significance, the relationship appears to be too weak to have educational significance.

Instructional Background

Interview subjects were asked, "Where did you learn about graphs?" This question was intended to define the instructional background of the subjects, particularly the content area of graph instruction.

PERCENT OF SUBJECTS REPORTING CONTENT SETTING OF GRAPH INSTRUCTION BY INTERVIEW GROUPS

	Low Scorers N=15	High Scorers N=15	Random Groups N=30
Mathematics	26.7	26.7	30.0
Study Skills	26.7	33.3	60.0
Reading	-	6.7	3.3
Language Arts	-	13.3	6.7
Social Studies	13.3	20.0	-
Don't Know	6.7	-	-
Non-Defined Answers	26.7	-	-

One category revealed may need definition, Study Skills is one of the three areas of the Wisconsin Design for Reading Skill Development for which instruction and record-keeping are mandated at the elementary level in the school district where this study took place. Although Study Skills is considered a part of the reading program objectives of the district, seventh graders did not make that distinction.

The emergence of the two categories, Mathematics and Study Skills, suggests that the instructional background of subjects was one that emphasized graphs isolated from a textual setting.

Feelings About Graphs

Subjects were asked the question, "Do you like graphs?" Responses were grouped into the three categories of Yes, No and Ambivalent. The Ambivalent category had to be created since a fair number of subjects appeared to hedge on answering this question. It was the generalized impression of this interviewer that subjects were somewhat concerned about the interviewer's status within the educational establishment: "If you are interviewing me about my answers and about graphs, you must like graphs?" Much of the ambivalence was expressed verbally with comments such as: "Sometimes yes, sometimes no" and "It depends." Non-verbal behavior included shrugs of the shoulders, eye contact and a searching of the interviewer's face for cues, or a refusal of eye contact accompanied by vague answers.

PERCENT OF SUBJECTS REPORTING POSITIVE, NEGATIVE, AND AMBIVALENT FEELINGS ABOUT GRAPHS BY INTERVIEW GROUPS

	Low Scorers N=15	High Scorers N=15	Random Groups N=30
No	46.7	6.7	30.0
Ambivalent	26.7	40.0	20.0
Yes	26.7	53.3	50.0

As can be seen, low scorers were more likely to say no, while high scorers tended to say yes.

Attention Factor

Subjects were asked two questions: "Do you ignore graphs when reading in a textbook?" and "When do you stop and pay attention to a graph?" The questions may appear to be positive and negative

ways of asking the same question, but the categories of responses to each question are somewhat different.

The most surprising finding was in the identification of the two categories, Teacher Directed and Text Directed.

Subjects described teacher direction in the form of classroom assignments, quizzes and tests, and direct instruction. Many subjects used such words as: "If the teacher tells me it is important, like it will be on a test, then I study the graph."

Some few subjects, generally high scorers, volunteered that they were paying attention to graphs as a result of direction in the textbook to "look at the graph" or were using study questions in the text which required extracting information from the graphic format.

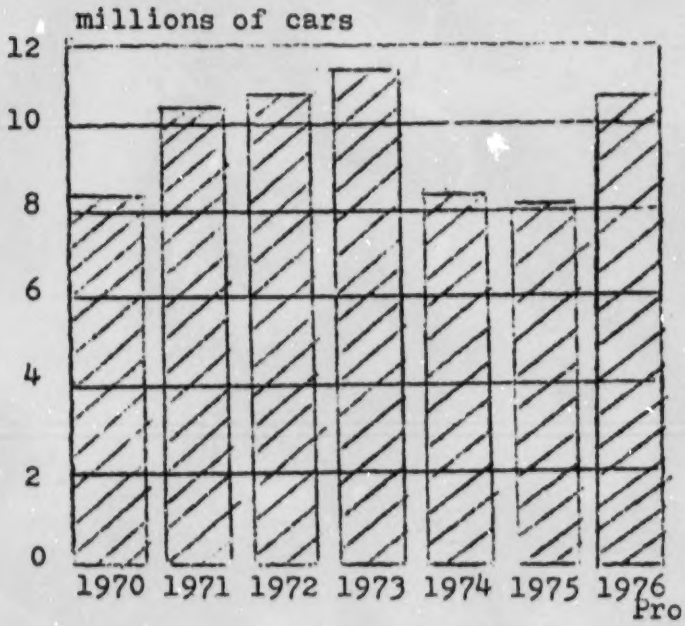
PERCENT OF SUBJECTS WHO SAY THEY IGNORE OR PAY ATTENTION TO GRAPHS BY INTERVIEW GROUPS

	Ignore			Pay Attention		
	Low N=15	High N=15	Random N=30	Low N=15	High N=15	Random N=30
Always	33.3	-	20.0	6.7	6.7	3.3
Sometimes	13.3	60.0	70.0	-	-	-
Teacher Directed	-	-	-	46.7	33.3	33.3
Text Directed	-	-	-	-	20.0	10.0
Never	40.0	40.0	6.7	-	-	-
Don't Know	13.3	-	-	-	33.3	20.0
Non-Defined Answers	-	-	3.3	46.7	6.7	33.3

Rating of Graph Difficulty

What happens when attention is focused upon the graphic presentation? What are the problems and/or distractions which bother the reader? To get at these questions, subjects were asked to identify the easiest and hardest of the eight graphs used in the investigation. These two questions elicited judgments about

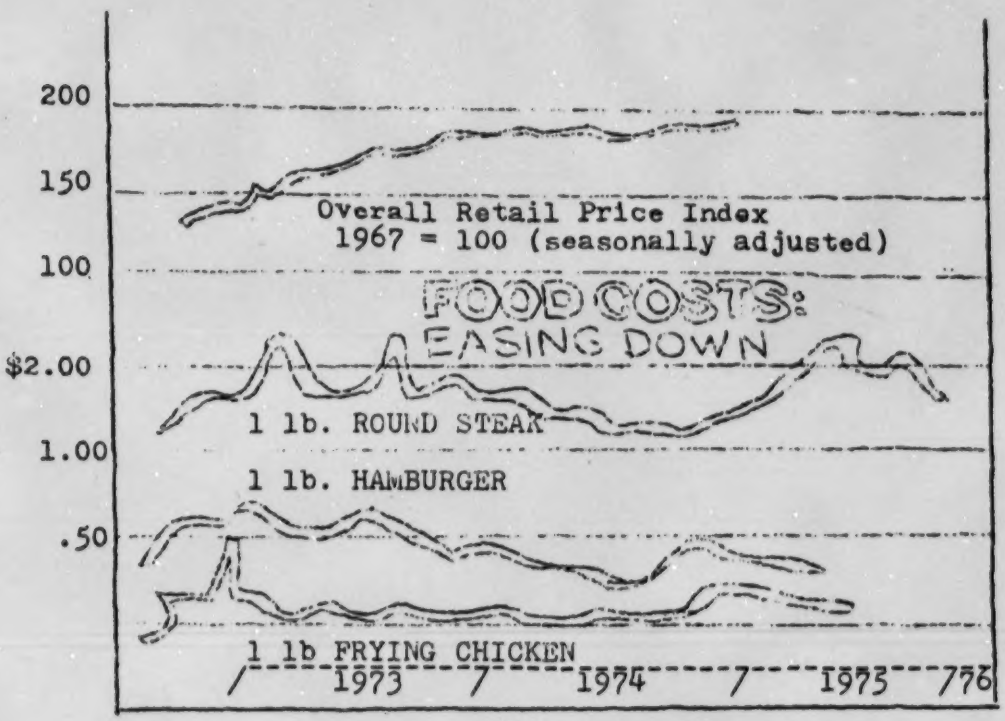
the readability and understandability of the graphic presentations and had little relationship to how subjects actually answered the multiple-choice items related to the graphs. The ranking achieved tended to support the logical principle that the more things that are compared and displayed on a graph, the more difficult the graph becomes. The principle can be affected, however, by a number of factors in the way the graph is drawn, the more noticeable being related to the labeling of the comparisons and additions to graphs (probably assumed by the illustrator to be "eye-catching.") Generally, seventh grade interview subjects saw graphs as difficult where the illustrator had tried to put "too much" into a compact space. Two of the graphs used in the investigation are reproduced to illustrate the perceptions of graph reading subjects.



U.S. NEW CAR SALES including imports

*Reprinted by permission from TIME, The Weekly Newsmagazine, Copyright Time, Inc., 1976.

Interview subjects saw this simple bar graph as the easiest and most understood graph. Factors contributing to this perception were the "flat-top" bars, clear vertical scale of twos, and non-interference from horizontal grid lines. A feature of the graph which did give some difficulty to subjects was the use of bold print for a part of the title, and less bold print and size for the words, "Including Imports." The abbreviation, "Proj.", printed at the base of the 1976 bar tended to be noticed by high scorers and not noticed by low scorers.



Interview subjects saw this complex line graph as the most difficult and one subject described the graph with the words, "It's going up and down all over it." Subjects talked about a number of features of the graph including calendar markers, four graphs in one, interference of grid line and information line labeling, and title information which ran counter to the trend of upward prices.

*Reprinted by permission from TIME, The Weekly Magazine, Copyright Time, Inc., 1976.

Item Analysis

Although an attempt had been made to define each item in the study in terms of what was expected of the reader (reading, interpretation, summarization), the interviews revealed that readers' reactions to each item were distinctive, eliciting behaviors which could be categorized. Two items which utilized the most difficult graph Food Costs: Easing Down, are reproduced. First is the graph reading item.

The cost of a pound of round steak in June of 1975 was:

a. almost $1
b. 60 cents
*c. over $2
d. $1.50

Sixty-two percent of 149 subjects answered c and thirty-one percent answered d. During interview it was revealed that the bulk of those answering d seemed to be paying attention to the words "round steak" in the stem of the item, and then located those words on the graph. The label, "1 lb. Round Steak" was extremely close to the $1.50 grid line, causing confusion of the grid line with the information line. By contrast, those who answered correctly (over $2) were paying attention to the date (1975) in the stem of the item. They located 1975 on the horizontal scale, moved upward to above the $2 grid line to the information line.

The graph summarization item which accompanied the graph also elicited interesting behavior.

The graph shows:

a. meat prices have stayed the same since January 1975
b. high meat prices appear to be going even higher
c. prices of three meats
*d. meat prices compared with the retail price index

The two most popular choices for low scorers were b and c with those answering c expressing much confidence in their answers and apparently not even looking at the top of the graph. The two most popular choices for high scorers were c and d. Again, as with low scorers, the high scorers answering c were confident that they were right, disregarding the top of the graph.

Conclusions and Implications

One of the most remarkable aspects of the study was in the fact that of the subjects interviewed, all could "read" and no interviews were terminated because of an inability to cope with the test. At the same time, these seventh graders were very similar to first graders learning to read. They exhibited many of the behaviors usually associated with beginning reading and/or frustration which included reading aloud, whispering, sub-vocalization, following lines of print with a forefinger, following grid and information lines with fingers, and relying on fingers as measurement devices. (Only two of 59 interview subjects demonstrated use of straight-edge measurement devices.)

In explaining their answers, subjects seemed to fall into four major groups, quite outside the designation of high or low scorers:

1. The first group did not have a strategy for achieving an answer saying, "I don't know" or "I guessed."

2. The second group exhibited attraction to, and distraction by, isolated words and phrases, reducing the task to what was known in much the same way a first grader will say "big" for "pig" because the word looks like one he knows. This group was quite susceptible to perceptions of size and height.

3. The third group, also susceptible to perceptions of size and height (bigger is better, higher is better), did have a strategy and expressed confidence in their answers. They exhibited a restricted, limited use of available information and did not confirm their answers.

4. The last group applied a strategy with accuracy, were direct and coherent in their explanations, and were likely to have a method for confirming answers.

The members of the above groups changed with each item in the study, evidenced by low scorers applying strategies with confidence on certain items and high scorers saying, "I don't know." Each individual then, might fall into any of the four groups described on any one item.

There was a stability, a repetitive "stuck" quality to the way that a particular subject dealt with a particular item. For example, if a subject gave an answer which, in the judgment of the interviewer indicated attraction and/or distraction by isolated words and phrases, and then wished to change an answer, the subject appeared to change to another answer for the same reason. The procedures in force during the interviews were intended to inhibit the investigator from introducing the instructional variable. The "aha" experience, the sudden dawning of insight which is so typical of first graders, was not seen and is cited as evidence of effective procedures (Roller, 1979, p. 84-149).

Recommendations for Graph Instruction

In combining the finding of the study related to (1) weak relationships with standardized achievement scores in reading comprehension and mathematics computation, (2) better scores on

graph reading and interpretation items when graphs were seen without a supporting text, and (3) a reported dependence upon the teacher to direct attention to graphs, two major recommendations for graph instruction emerged.

The first recommendation is to the content teacher: whenever a graph is encountered in a textbook, the opportunity should be seized to improve and reinforce the graph skills of students. An excellent reference for teachers is found in Summers (1965). In addition, listed below are a number of questions which are intended to aid the planning of the teacher when graphs are encountered in a textbook and may be suitable as a guide for construction of an informal inventory of students' graph skills.

Purpose: Why did the author put a graph at this point in the text? What is the author trying to say that can be shown only (or best) in a graph? What is the relationship of concepts developed in the text to concepts shown on the graph? Does the author expect the reader to make the connections between text and graph or is some help provided?

Readability and Understandability Factors: What kind of graph is pictured? How many relationships are pictured? Is it a simple or a complex graph? What are the factors in the way the graph is drawn which may prove difficult for the reader? Is it printed in color? What is the placement of titles and labels? Will students see this graph as easy or difficult, and why?

Titles and Labels: What comparisons are made on the graph? How are those comparisons labeled? What does the title say? Is the title complete or incomplete? Are different types of print used in labels and title; if so, is there any significance?

Measurement Scale: What is the unit of measurement? Are grid lines used? Do grid lines help or hinder the reader? Does the scale begin at zero? What is the highest point on the scale?

Time Units: What unit of time is used: days, months, years? Are any unusual conventions used? Are any abbreviations used?

Symbols, Keys, Legends: Why is a legend important to this graph? Is the legend clear? Where is it placed?

Information-bearing Areas, Lines, Bars, Pictures: What information should the reader glean from this graph? Is the reader expected to read directly stated information? Is the reader expected to make interpretations? Is computation necessary? Which operations of addition, subtraction, multiplication and/or division are necessary? How can students be encouraged to confirm their answers?

Trends: Is a trend pictured on the graph? How complete is the information? Has bias been introduced through the picturing of incomplete information? Are there conflicting pieces of information in the titles or labels?

Integration with Text: What major pieces of information contained in the graph are discussed in the text? What information is not discussed? Are there any contradictions between graph information and text information?

Evaluation of Presentation: Did the author accomplish his purpose by putting a graph at this point in the text? Was the graph understandable? How might the presentation have been improved? How important is the information on the graph to class and course requirements, and teacher expectations?

The second recommendation is for the content teacher who must manage the classroom, complete with the highs and lows of student abilities. It is likely that any student at any given time may have difficulty with a given graph because each graph, whether in a textbook, periodical, or reference book, is unique in terms of the purposes of the author and illustrator. Therefore, instruction focusing on the graph will seldom be repetitive or

redundant. Any major resistance noted will probably come from the student who is totally lost and has no strategy for extracting meaning. When this is the case, an opportunity for direct instruction is available.

For the on-looker student when direct instruction is occurring, opportunity is available to reinforce the strategies which are most efficient in extracting information from the graphic format. If there is a problem with the on-looker, it will probably take the form of the student who wants to blurt out the answer, thereby interferring with the necessary thinking time of other students.

One method recommended to cope with this situation would be the use of written questions which each student would answer for himself. However, an essential second step would include allowing students to interact about how their answers were obtained, preferably in the form of student-to-student interaction rather than student-to-teacher. If students were given the task of resolving any differences in answers in small groups, the teacher could rotate among the groups, focusing attention upon ignored information, and encouraging an ever expanded view of information contained in the graphic format.

References

Buswell, G. T. How adults read. Supplementary Educational Monographs. Chicago: Univ. of Chicago Press, 1937.

Eggen, P., Kauchak, D., & Kirk, S. The effects of generalizations as cues on the learning of information from graphs. Journal of Educational Research, April/May, 1978, pp. 211-213.

Food Costs: Easing Down. Time, April 19, 1976, p. 71.

McCallister, J. M. Reading difficulties in studying content subjects. Elementary School Journal, November, 1930, 31, 191-201.

Otto, W., Kamm, K., & Weible, E. Wisconsin design for reading skill development: rationale and objectives for the study skills element, working paper, No. 84. Madison: Wisconsin Research and Development Center for Cognitive Learning, February, 1972, pp. iii; 13-14; Appendix.

Roller, Beverly Van Orden. Graph Reading Abilities of Seventh Grade Students. Unpublished doctoral dissertation, State University of New York at Buffalo, 1979.

Shores, J. H. Skills related to the ability to read history and science. Journal of Educational Research, April, 1943, 36, 584-593.

Strickland, R. G. A Study of the possibilities of graphs as a means of instruction in the first four grades of the elementary school. New York: Teachers College, Columbia University Contributions to Education, No. 749, 1938.

Summers, E. G. Utilizing visual aids in reading materials for effective learning. In Herber, H. L. (Ed.), Developing study skills in secondary schools. Newark: International Reading Association, 1965, pp. 97-155.

Turner, B. The extent to which selected sixth grade student's use of graphic aids enhances their comprehension of content material. Unpublished doctoral dissertation, University of Northern Colorado, 1974.

U. S. new car sales including imports. Time, June 14, 1976, p. 54.

Vernon, M. D. Learning from graphical material. British Journal of Psychology, 1946, 36, 145-158.

Vernon, M. D. The visual presentation of factual data. British Journal of Educational Psychology, November, 1950, 20, 174-185.

Vernon, M. D. Learning and understanding. Quarterly Journal of Experimental Psychology. February, 1951, 3, 19-23.

Vernon, M. D. The value of pictorial illustration. British Journal of Educational Psychology, November, 1953, 23, 180-187.

Vernon, M. D. The instruction of children by pictorial illustration. British Journal of Educational Psychology, November, 1954, 24, 171-179.

A Comparison of Three Editions and Sample Frequency of the Dale-Chall Readability Formula

Robert T. Williams, Ed.D
Department of Education
Colorado State University

Teachers are concerned with helping students learn by assigning them material which is as readable or learnable as possible. Teachers vary widely in their ability to accurately judge the difficulty of materials; their assessments are based on the ability levels of the children they teach and a common sense of grade level does not exist (Jorgenson, 1975, Miller, 1979). Teachers who wish to examine the relative difficulty between and within instructional materials seem to need to use a readability formula. Readability estimates have been criticized from at least two sides; the measures consider a limited number of factors in assessing or estimating readability and users have attempted to match material readability scores and student standardized test scores. But, given these limitations, which will be considered in more detail below, a readability formula will yield a quick, if perhaps dirty, estimate of the relative difficulty of instructional material. Blair (1971) asserts that readability formulas give "probability statements of student

---

[1]This research was completed while the author was on sabbatical leave at the University of Kansas, Lawrence, Kansas and the Reading/Writing in the Content Areas Project at the University of Northern Colorado, Greeley, Colorado. The author expresses thanks to Dr. Dick Tracy, School of Education, University of Kansas, for assistance in the statistical analysis and to Dr. Robert A. Pavlik and Dr. Beverly Roller, the University of Northern Colorado, for their insightful suggestions in the preparation of this manuscript.

success." Furthermore, Guthrie (1972), using 62 sixth grade students in a Catholic school in Baltimore, concludes that readability approximates learnability. He summarizes his study by saying,

> Thus, the same variables which are contained in many readability formulas, including the Dale-Chall index, predict with considerable accuracy the amount of new learning which occurs as a result of reading. Materials which are judged to have high readability for a given age group (grade level) will likely impart a substantial amount of learning to students of that age (grade) who read the materials (p. 279).

This study compares and contrasts three editions of the Dale-Chall Readability Formula and explores the effect of sampling text material less frequently than directed by formula instructions. The first part of the study examines the relationship between the 1948 (Dale and Chall, 1948), 1958 (Powers, Sumner and Kearl, 1958) and the 1968 (Holmquist, 1968) editions of the Dale-Chall Readability Formula. The 1948 Dale-Chall used the 1926 edition of the McCall-Crabbs Standard Test Lessons in Reading (McCall and Crabbs, 1926) as a criterion test. Powers, Sumner and Kearl (1958) observed significant differences in vocabulary and subject matter between the 1926 and 1950 editions of the McCall-Crabbs Standard Test Lessons in Reading (McCall and Crabbs, 1950). Holmquist (1968) used the 1961 (McCall and Crabbs, 1961) edition of the test lessons in his recalculation. The equations used for each edition of the Dale-Chall are:

1948: $X_{c50} = .9496$ (sentence length) + .1579 (Dale Score) + 3.6365

1958: $X_{c50} = .0596$ (sentence length) + .1155 (Dale Score) + 3.2672

1968: $X_{c50} = .0596$ (sentence length) + .1142 (% of non-Dale Words /Dale Scores/) + 3.4442*

*Holmquist also attempted to supplement the Dale List of 3,000 Familiar Words with 102 "familiar" science words. However, there is no practically significant difference, probably no statistically significant difference, when these 102 "familiar" science words are added to the Dale list.

where $^{x}c50$ is the reading grade a student must have attained to understand 50 percent of the analyzed passage; sentence length is the average sentence length in the sample: Dale Score is the percentage of unfamiliar words (words not on the Dale list of 3000 Familiar Words) in the sample.

Studies

To examine the relationship between these three editions of the Dale-Chall Readability Formula, twenty-three texts used in the public school were examined. Eight social studies texts, ten science texts, two composition texts and one text from mathematics, one text from business and one text from health care were used. Samples of from 100 to 200 words were taken from every tenth page of the texts. Formula raw scores were computed for each edition of the Dale-Chall. The formula raw scores were correlated to determine the relationship between the editions of the Dale-Chall formula and a single-factor ANOVA with repeated measures (Weiner, 1971, 267-69) was used to test for significance between the three editions.

The tables that follow present the data used in this study. Table I gives the mean formula raw scores using each edition of the Dale-Chall with the 23 texts. Table II shows the correlations between the three editions. Table III shows the ANOVA between the three editions of the Dale-Chall.

TABLE I
MEAN, FORMULA RAW SCORES FOR TWENTY THREE TEXTBOOKS USING THE 1948, 1958, and 1968 EDITIONS OF THE DALE-CHALL READABILITY FORMULA

Text	No. of Samples	Formula	Formula	Formula	Total
T 1	80	8.12	7.03	6.99	22.14
T 2	15	7.40	6.49	6.47	20.36
T 3	26	6.77	5.99	5.99	18.75
T 4	62	6.58	5.79	5.81	18.18
T 5	70	7.93	6.86	6.86	21.65
T 6	42	8.92	7.78	7.69	24.39
T 7	44	9.29	7.96	7.88	25.13
T 8	39	8.09	6.92	6.92	21.93
T 9	30	6.40	5.55	5.62	17.57
T 10	20	8.48	7.28	7.26	23.03
T 11	34	6.52	5.64	5.73	17.89
T 12	44	7.81	6.79	6.77	21.37
T 13	24	5.55	4.97	4.86	15.38
T 14	29	5.99	5.25	5.32	16.57
T 15	25	7.54	6.52	6.53	20.59
T 16	24	8.34	7.21	7.17	22.72
T 17	21	7.23	6.33	6.33	19.89
T 18	25	9.07	7.71	7.67	24.45
T 19	48	7.26	6.35	6.33	19.94
T 20	15	8.79	7.53	7.49	23.81
T 21	14	9.24	7.78	7.78	24.80
T 22	73	5.82	5.18	5.23	16.23
T 23	47	7.60	6.49	6.55	20.64
		174.75	151.40	151.25	477.4=G
		F_1	F_2	F_3	
		$\bar{X}_{F1}$=7.5978	$\bar{X}_{F2}$=6.5826	$\bar{X}_{F3}$=6.5761	

TABLE II
CORRELATIONS BETWEEN THE 1948, 1958, and 1968 EDITIONS
OF THE DALE-CHALL READABILITY FORMULA

	1948	1958	1967
1948	--	.9977*	.9986*
1958		--	.9987*

* $p < .01$

TABLE III
SINGLE FACTOR ANOVA WITH REPEATED MEASURES
FOR THE 1948, 1958, 1968 EDITIONS
OF THE DALE-CHALL READABILITY FORMULA

Source of Variation	SS	df	MS	F
Between texts	60.8880	22		
Within texts	16.7637	46		
Formula	15.9057	2	7.9529	407.8410*
Residual	.8580	44	.0195	
Total				

*$p < .01$

The significant F value in Table III indicates a statistically significant difference among the three editions of the Dale-Chall. However, the source of the difference is not readily apparent. Inspection of Table I shows that the 1948 edition is associated with the highest mean formula raw scores. Therefore, the hypotheses that the 1948 edition is significantly different from the 1958 and the 1968 edition was tested following Meiner (1971, 167-69). An F value of 815.6531 ($P < .05$) shows that the 1948 edition is significantly different from the 1958 and the 1968 editions. The hypothesis that the 1958 and 1968 editions were significantly different from one another was not rejected ($F = .025$, non significant, Scheffe $F^{.05}_{1,40}=8.16$).

The high, positive, significant correlations seen in Table II suggest that all three editions of the Dale-Chall are measuring the same factors. However, there is a statistically significant difference between the 1948 edition and its later revisions. The 1948 formula evaluates material as being significantly more difficult than evaluated by the 1958 or the 1968 edition. There is no significant difference between the 1958 and the 1968 revisions.

It seems, then, that the changes from the 1926 to the 1950 editions of the McCall-Crabbs Standard Test Lessons in Reading observed by Powers, Sumner, and Kearl are reflected in their formula revision and that the changes in the 1961 McCall-Crabbs lessons observed by Holmquist are reflected in his formula revision, but there is no significant difference between the Powers, Sumner and Kearl (1958) and the Holmquist (1968) revisions.

Since there is a table for rapid determination of the Dale-Chall as revised by Powers, Sumner and Kearl (Williams, 1972), and since the 1958 revision reflects changes in the criterion measure, the author suggests using the 1958 revision of the Dale-Chall.

Whichever version of the Dale-Chall one might select, the application of the formula still is time consuming, even with shortcut measures like those developed by Kearl (1952), Koenke (1971) and Williams (1972). The second part of this study examines the effect of sampling text material less frequently than directed by the formula.

If equivalent results could be achieved by sampling the materials less frequently than the every 10 pages dictated by the Dale-Chall instructions, much time could be saved. Burkhead and Ulferts (1977) examined 48 introductory college management textbooks using a computer application of the Dale-Chall Readability Formula to determine if there was a significant difference in sampling every 10th, 20th, 30th, 40th, or 50th page of the text. Using a t-test, they found only one text at 40-page intervals to be significantly different at .05 level of confidence. They concluded that:

> Samples taken from 48 introductory college management textbooks at 20-page, 30-page, and at 50-page intervals yield as dependable Dale-Chall scores as those taken at 10-page intervals. (288)

They emphasize that "In 47 of the 48 textbooks, results taken at 40-page intervals were just as dependable as those taken at 10-page intervals..." (288) Martin and Lee (1961) found no significant difference between the means obtained when sampling every 10th and every 50th page in high school biology textbooks.

Discussion

These studies provide an interesting direction for thinking about sampling in readability analysis, but both studies limited their evaluations to one content area (Burkhead and Ulferts to management at college level and Martin and Lee to biology texts in high school). To provide a wider perspective to this question of sampling, the textbooks described above were analyzed for

differences in mean formula raw scores of samples taken every 10th, 20th, 30th, 40th and 50th pages. Both the 1948 and 1958 versions of the Dale-Chall Readability formula were used. Since no significant differences were found between the 1958 and the 1968 versions as reported above, the 1968 version was not analyzed. The texts were sampled every 10th page with passages of 100 to 200 words. The formula raw scores were calculated by hand calculator following formula instructions, using the worksheet and the Kearl table (Kearl, 1952) for the 1948 version and the Williams table (Williams, 1972) for the 1958 edition. The texts were then sampled at 20-, 30-, 40-, 50-page intervals by examining the worksheets for each text.

Each text was tested for significant differences between the 10-, 20-, 30-, 40-, and 50-page interval samples using ANOVA. Since twenty-three analysis of variance were calculated for each of the two editions of the formula, the F tables are not presented here. The raw data and F tables may be obtained from the author. There was no significant F value found for any of the twenty-three texts using either the 1948 or the 1958 edition of the Dale-Chall. These results suggest that one would not obtain significantly different results from a Dale-Chall readability analysis by sampling the text every 10, 20, 30, 40, or 50 pages. Therefore, the amount of time required to use the Dale-Chall formula may be reduced by at least half.

For teachers who wish to examine the relative difficulty between or within instructional materials, using the 1948 or 1958 edition of the Dale-Chall and sampling texts every 20-, 30-, 40-, or 50-pages seems appropriate. But the reader should be cautioned against attempting a one-for-one match between a textbook's readability score and a reader's standardized reading score. In attempting to determine the appropriateness of a book for a child or group of children, one might be guided by Ekwall and

and Henry (1968), Thelen (1976, p. 6-11), Earle (1976, p. 61-68), Williams and Kaman (1975) or Lundstrum and Taylor (1978, p. 10-20). Furthermore, users of readability formulas are reminded that the Dale-Chall, and most other formulas, consider the two factors of vocabulary and sentence length in the measurement of readability. Other factors are being suggested and explored to refine the measurement of readability.

Endicott (1973) and Granowsky and Botel (1974) have suggested procedures for examining the syntactic complexity of written material as a measure of its difficulty. Other researchers have explored the relationship of language functions and readability. Pavlik (1977) suggests that textbook language serves to (1) control action - reading material such as directions, manuals, formulas, handbooks, and indexes have the potential to control the reader's action; (2) order thinking - reading material such as definitions, laws, background explanations, arguments, introductions and summaries have the potential to order the reader's thinking; (3) promote abstract thinking - reading material such as predictions, hypotheses and theories have the potential to promote the reader's abstract thinking. McMillin (1977) found that the mean readability estimates of biology material differed significantly among passages with language to control action, language to order thinking and language to promote abstract thinking. McMillin concluded that "These differences indicate the function of language may be a significant factor in the readability of a passage..." Further, using narrative and expository passages closely controlled for readability, Brazee (1976 and this monograph) found that eighth grade students' reading levels were higher on narrative passages than on expository passages when reading levels were measured by a reading miscue inventory and an informal reading inventory. The type of writing seem to influence students' ability to read the material.

Another direction of questioning and research might be the interaction of the reader's level of cognitive development, the cognitive level of the concepts in the instructional material, and the cognitive level of the language used to carry the concepts, that is, a learner X concept X language interaction. Also of interest might be an examination of the relative difficulty of words, or the same word, used in different context, carrying different levels of meaning, or having different Rogetian categories of meaning (Chapman, 1977, xvii - xxiv). Beyond an estimate of relative difficulty, the utility of current readability measures is limited; questions to refine the measurement of readability are evolving; the prospects for future research are exciting.

References

Blair, Allen M., "Everything You Always Wanted to Know About Readability Formulas but were Afraid to Ask," English Journal, 12, winter, 1978, 67-75.

Brazee, Phyllis E., "A Qualitative and Quantitative Description of Eighth Grade Students' Oral Reading in Both Narrative and Expository Material," Ed.D. Dissertation, University of Northern Colorado, Greeley, Colorado, 1976.

Burkhead, Marie and Ulferts, Greg "Sample Frequency in Application of Dale-Chall Readability Formula," Journal of Reading Behavior, 9, 3, Winter, 1977, 287-90.

Chapman, Robert L. (Revisor) Roget's International Thesaurus, Fourth Edition, Thomas Y. Crowell, New York, 1977.

Dale, Edgar and Chall, Jeanne S. "A Formula for Predicting Readability: Instructions," Educational Research Bulletin, 27, 1948, 37-54.

Earle, Richard A. Teaching Reading and Mathematics, International Reading Association, Newark, Delaware, 1976.

Ekwall, Eldon E. and Henry, Ida B. "How to Find Books Children Can Read," The Reading Teacher, 22, 1968, 230-32.

Endicott, A. L. "A Proposed Scale for Syntactic Complexity," Research in the Teaching of English, 7, 1, 1973, 5-12.

Granowsky, Alvin and Botel, Morton "Background for a New Syntactic Complexity Formula," Reading Teacher, 28, 1, 1974, 31-35.

Guthrie, John T. "Learnability Versus Readability of Texts," The Journal of Educational Research, 65, 6, February, 1972, 273-280.

Holmquist, John B. "A Determination of Whether the Dale-Chall Readability Formula May be Revised to Evaluate More Validly the Readability of High School Science Materials," unpublished doctoral dissertation, Colorado State College, Greeley, Colorado, 1968.

Jorgenson, Gerald W. "An Analysis of Teacher Judgements of Reading Level," American Educational Research Journal, 12, Winter, 1975, 67-75.

Klare, George R. "A Table for Rapid Determination of Dale-Chall Readability Scores," Educational Research Bulletin, 31, 1952, 43-47.

Koenke, Karl "Another Practical Note on Readability Formulas," Journal of Reading, 15, 1971, 203-08.

Martin, W. and Lee, W. "Sample Frequency of Application of Dale-Chall Readability Formula," Educational Research Bulletin, 40, September 13, 1971, 146-49.

McCall, W. A. and Crabbs, L. M. Standard Test Lessons in Reading, Columbia University Teachers College, New York, 1926, 1950, 1961.

McMillin, Larry M., Allen, Ilene, Clement, Betty and Brown, Carroll "Language Complexities not Measured by Fry and Fog Readability Formulas," Unpublished study, University of Northern Colorado, 1977.

Miller, John W. "Teachers' Abilities to Judge the Difficulty of Reading Materials," Reading Horizons, 19, 2, Winter 1979, 151-58.

Pavlik, Robert A. Reaching for Reading Maturity in the Content Fields, (A coursebook to accompany EDRD 521, Reading in the Content Fields), University of Northern Colorado, Greeley, Colorado, 1977.

Powers, R. B., Sumner, W. A., and Kearl, B. E. "A Recalculation of Four Readability Formulas," Journal of Educational Psychology, 49, 1958, 99-105.

Thelen, Judity Improving Reading in Science, International Reading Association, Newark, Delaware, 1976.

Weiner, B. J. Statistical Principles in Experimental Design, McGraw-Hill Book Company, New York, 1971, 267-79.

Williams, Robert T. "A Table for Rapid Determination of Revised Dale-Chall Readability Scores," The Reading Teacher, 26, 1972, 158-65.

Williams, Robert T. and Kaman, Mitchell, A., "A Model for the Development of a Reading and Study Skills Inventory for Every Content Area," in Sugimoto, Roy (Ed.) College Learning Skills: Today and Tomorrowland, Proceedings of the English Annual Conference of the Western College Reading Association, 1975, 211-16.

Study Guides as Aids to Learning: A Review of Research*

Jodi Grant
Reading Department
University of Arizona

Because much of the material written for use at the junior high and high school level is such that students have difficulty comprehending it without guidance, study guides have been proposed as one means of providing such guidance. Study guides may be differentiated from worksheets in that study guides are usually developed along the lines of a clearly identifiable theoretical base. It cannot be assumed that providing a student with written material will result in learning; rather, students must be taught how to learn (Herber, 1970).

The value of study guides is based on the notion that if material is too difficult or if the concept load is heavy, guides can be employed to aid students' understanding (Estes and Vaughan, 1978; Herber, 1978; Honeycutt, 1973; Tutolo, 1977). Guides are generally described as intended to give direction to the use of appropriate learning strategies and further to aid in the understanding of the specific content information; thus, content understanding and reading ability can be developed simultaneously (Herber, 1978). Guide material may additionally aid the student in perceiving organization and in focusing on structure (Vacca,

---

*Ed. Note: This review is based on available studies primarily conducted at the Reading and Language Arts Center, Syracuse University, Syracuse, N.Y. Other studies that may have been completed and that are not included in this review may be obtained from Dr. Hal Herber at Syracuse University.

1973). Although referring to study guides as "crutches," Estes and Piercey (1973) go on to suggest that such guidance can aid the student who must use material that would otherwise be beyond his/her grasp. Ultimately, such "crutches" lead students to becoming independent learners.

The format of study guides varies. Herber (1967), who is often credited with the development of study guides and who certainly has given them widespread exposure, provided the following general description of study guides:

> These guides help the students apply skills in such a way that they are conscious of the process involved in the skill and are also aware of the concepts being developed and applied. (They) are so constructed that, in one class, students can react to the same material at different levels of comprehension, thereby serving the levels of ability and achievement represented in the class. (Study guides are) designed to serve as the basis for reading and discussion of the major concepts of the unit being studied (p. 101).

Briefly, the literal level study guide is designed to deal with what the author said; the interpretive level study guide is designed to aid the student in deriving meaning from what the author has meant; the applied level study guide takes the reader beyond both the literal and interpretive levels in that the reader applies what the author has said in some pragmatic or theoretical exercise (Herber, 1970). In his book, Teaching Reading in Content Areas, Herber (1970; 1978) extensively describes the three levels of comprehension and the rationale for using study guides based on these three levels of understanding. He further provides a wealth of sample study guides to help teachers understand how to construct them.

Purpose

Study guides are often suggested as a means by which the content area teacher can assist students when they are required

to read material that may prove difficult for them (Earle, 1969; Estes and Vaughan, 1978; Hafner, 1977; Herber, 1970, 1978; and Tutolo, 1977). The purpose of this paper is to review the research that has been conducted on the use of study guides to determine if these guides have been shown to be of benefit to junior high and high school students' comprehension and/or reading ability.

Review of the Research

A study by Ellsworth Berget (1973) investigated two methods of guiding students in the reading of a short story: (1) study guides and (2) purpose-setting questions. Included in the study was a control group for comparison purposes. The sample for the study consisted of 81 ninth-grade students chosen from five English classes in a "working class suburb" of Syracuse, New York. With the exclusion of educationally handicapped and mentally retarded, all ability levels were represented. Students in each class were randomly assigned to one of the three groups.

Experimental group I used three-level study guides with the reading selection. This group worked in a classroom with the investigator who explained the purpose and use of the study guides to the students before they were asked to read the selection. Students completed level one guides while reading or after having completed the reading of the selection. Answers to the study guide were discussed in small groups without teacher guidance following the completion of the reading.

Experimental group II were given purpose-setting questions with the reading selection. They remained in their original classroom with their English teacher. Before they read the selection, the purpose and use of the purpose setting questions were explained by the teacher. Small groups were set up after the reading in order to discuss and answer the questions.

Experimental group III, the control group, was sent to the library where they read the selection under the supervision of the librarian. To approximate classroom conditions, no one else was allowed in the library.

A post-test only control group design was utilized and a one-way analysis of variance was used to analyze the data. An experimenter-constructed test, consisting of 25 multiple choice items reflecting three of the levels of understanding as delineated by Bloom (1956) (knowledge, comprehension, application) was used as a measure of comprehension.

The results indicated a significant difference in favor of those students who received the study guides. The differences were significant at the .05 level between study guides and the control group as well as between study guides and purpose setting questions. There were no significant differences between purpose setting questions and the control group. The investigator concluded that the use of written study guides enhances learning to a greater extent than do purpose setting questions or simple instructions to read the passage.

In a pilot study conducted by Estes (1969), ninth grade social studies students were divided into three groups. One group used study guides plus small group discussion; a second used study guides alone; and a third group used purpose-setting questions. Three criteria were measured: (1) gain in general reading comprehension (using the Stanford Reading Test), (2) change in attitude toward reading, and (3) social studies achievement. Results of the study were considered inconclusive, as the probability levels associated with treatment differences did not reach the required levels of significance (.05). On the basis of this study, several revisions were made and a more thorough study was designed.

The subsequent study conducted by Estes (1973) was divided into two parts. The first part of the study compared students who received (1) guide material with small group discussion, (2) guide material with no small group discussion, and (3) no guidance or discussion. These variables were examined for their effect on (1) social studies knowledge, (2) attitudes, and (3) reading comprehension. The study guide was designed based on the literal, interpretive, and applied levels as delineated by Herber (1970).

Students in the ninth grade social studies classes were instructed by two teachers over a 15 week period. Each teacher taught one class using guide material and small group discussion. Another class was taught by each teacher using guide material and large group discussion. A third class (the control group) was given questions designed to facilitate retention of the material as proposed by Weaver and Kingston (1967). These questions were to accompany their reading. Specific social studies concepts served as a basis for both the guide material and the questions.

At the end of the instructional period, three measurements were administered. Social studies achievement was measured using an experimenter constructed test of social studies knowledge (reliability .86). An experimenter constructed scale was used to measure attitude toward social studies (reliability .87). Reading achievement was measured using the Stanford High School Reading Test. Planned comparisons were used to analyze the results which indicated that differences for reading comprehension favored the use of guide material with large group discussion (.05). Differences between guides and the control group were not significant. Results for social studies knowledge were the same as for reading comprehension. For the social studies attitude variable, significant (.05) differences favored the control group as opposed to the group using guide material.

The author drew three conclusions from this portion of the study. With regard to social studies knowledge and reading comprehension, there appeared to be no marked difference between guide material and prereading questions. Further, the use of small group discussion in the classroom appears to be of questionable value. The third conclusion was that attitude toward social studies cannot be expected to be affected in the same manner as the skill variables (knowledge and comprehension).

Part two of the study grew from questions that arose from part one. The teachers' use of three different procedures might have had an effect on the results. Since there was an interaction between teachers and procedures in the first part of the study, it was decided to assign each teacher to the procedure he had seemed to use most successfully in the earlier part of the study. In order to gain insight as to whether the assistance of guides or prereading questions was more or less valuable for poorer readers than for the whole group, the readers were assigned to two groups based on scores on the Stanford High School Reading Test. Students who scored below average (below a standard score of 50) were classified as the poorer readers; the remainder of the sample formed the other group.

Thus, for the second part of the study, one of the two teachers used guide material while the other used prereading questions. The original sample of subjects was used with students being grouped based on the previously mentioned criteria. Six weeks of instruction ensued. At the conclusion of this instruction, the students were administered an experimenter constructed test of social studies knowledge (reliability .76). Differences between guides and prereading questions for the entire group were not found to be statistically significant. In the case of the poorer readers, the difference was statistically significant (.05) in that prereading questions seemed to be more helpful than did guide material.

Based on these findings, the author drew three conclusions. First, guide material should be written in such a way that it does not make the task more complicated than it should be, especially in the case of poorer readers. That is, guide material should be written on a level that equals the students' ability. Second, since questions appear to be an effective strategy, the three levels of comprehension may be incorporated into these questions. Third, different types of guide materials may be necessary for the various reading abilities that may be found in schools.

In a study conducted by Sanders (1969), the researcher attempted to adjust instruction to the reading needs of students. Ninth grade social students in two classes were taught by the same teacher. One group was taught by means of lecture, assigned reading and follow-up or review. A more inductive approach was used with the second group in that pertinent background was established, purposes were given for reading assignments, vocabulary was pre-taught, assigned reading was accompanied by study guides, and small group activities were included to promote student participation and an exchange of views.

To evaluate the results, a pre- and post-test comparison of scores and a teacher assessment of pupil understanding and reaction were used. Pre-test scores revealed virtually no difference between the two groups; post-test scores showed a marked difference in favor of the group using the study guides. As regards to the teacher assessment of pupil understanding and attitude, the teacher rated the group using study guides as having a greater depth of understanding and as demonstrating a greater level of participation and enthsiasm than the control group. Sanders' conclusions centered on the belief that the use of guide material, purpose setting, etc., was superior to the teaching method wherein lecture and assigned reading were the strategies employed.

One criticism of this study centers around teacher bias. It was stated by the author that the teacher in question had been searching for a more effective teaching method and was highly pleased with the study guide format. No reference is made to the source of the test used for the evaluation and no level of reliability or validity is given. Only lower track students were used in the sample and only two classes formed the comparison groups. The study guide variable is confounded by other variables such as vocabulary instruction or small group activities.

Conclusions

Although research dealing with the use of study guides is far from extensive, some conclusions may be drawn from the research that has been conducted. First, the use of study guides does appear to be an aid to comprehension when used in conjunction with group discussion. Second, and perhaps most important, in order for the students to make efficient use of the guide, it should be written on a level commensurate with the students' abilities. Third, guides may function as a pre-reading, purpose-setting method and as a post-reading method.

Possibilities for fruitful research in the area of study guides is apparent. For instance, investigations as to the effectiveness of study guides as opposed to worksheets may prove revealing. Also, research exploring the pre-existing schema of the reader relative to the usefulness of study guides is in order. Such research would be designed to determine if the guides are beyond the students or if the students are beyond the guides. Research needs to be expanded to include all areas of curricula where guides may prove beneficial. In addition, research with various grade levels would be appropriate. Guides come in various formats, varying from three level (Herber, 1970, 1978) to cause-effect to almost anything the teacher wants

to make them. Estes and Vaughan (1978) discuss concept study guides that are based on the notion of chunking. The effect of various types of guides on learning could be a possible avenue for investigation. As stated by Samuel Weintrab (1971), "Compared with accomplishments in most other sciences, reading research is still in its infancy. A great many questions remain unanswered or, at best, only partially answered." (p. 178)

References

Berget, Ellsworth. Two Methods of Guiding the Learning of a Short Story. In H. L. Herber and P. L. Sanders (Eds.) Research in Reading in the Content Areas: Second Year Report. Syracuse: Syracuse University Press, 1973.

Bloom, B. S. et al. Taxonomy of Educational Objectives: Cognitive Domain. New York: David McKay and Co., Inc. 1956.

Campbell, Donald T. and Stanley, Julian C. Experimental and Quasi Experimental Designs for Research. Chicago: Rand McNally and Co., 1963.

Estes, Thomas H. Use of Prepared Guide Material and Small Group Discussion in Reading Ninth Grade Social Studies Assignments: Pilot Study Report. In H. L. Herber and P. L. Sanders (Eds.) Research in Reading in the Content Areas: First Year Report. Syracuse: Syracuse University Press, 1969.

Estes, Thomas H. Guiding Reading in Social Studies. In H. L. Herber and P. L. Sanders (Eds.) Research in Reading in the Content Areas: Second Year Report. Syracuse: Syracuse University Press, 1973.

Estes, Thomas H. and Piercey, Dorothy. Secondary Reading Requirements: Report on the State. Journal of Reading, 1973, 17, 20-24.

Estes, Thomas H. and Vaughan, Joseph L. Reading and Learning in the Content Classroom: Diagnostic and Instructional Strategies. Boston: Allyn and Bacon Company, Inc., in press.

Gardner, E. F., Merwin, J. C., Callis, R., and Madden, R. Stanford Achievement Test: High School Battery. New York: Harcourt, Brace and World, 1965.

Grant, Jodi and Keetch, Linda. The Golden State. Reading and Prospecting for Futures: Proceedings from the Tenth Annual Conference, Newport Beach, California: California Reading Association, 1976, 1, 136-138.

Hafner, Lawrence. Developmental Reading in Middle and Secondary Schools: Foundations, Strategies, and Skills for Teaching. New York: Macmillan Publishing Co., 1977.

Herber, H. L. Developing Reading Skills Demanded by Content Subjects. In Mildred Dawson (compiler) Developing High School Reading Programs. Newark, Delaware: International Reading Association, 1967, 99-102.

Herber, H. L. Teaching Reading in Content Areas. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1970.

Herber, H. L. Teaching Reading in Content Areas (2nd Ed.). Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1978.

Honeycutt, David. A Procedure to Investigate Constructs. In H. L. Herber and P. L. Sanders (Eds.) Research in Reading in the Content Areas: Second Year Report. Syracuse: Syracuse University Press, 1973.

Robinson. H. A. Corrective Reading in the High School Classroom: Some Principles and Procedures. Corrective Reading in the High School Classroom: Perspectives in Reading, 6, 1966, 11-22.

Sanders, Peter L. Teaching Map Reading Skills in Grade 9. Journal of Reading, 1969, 12, 283-286.

Singer, Harry. Reflections and Investigations on Reading. Twenty-fifth Yearbook of the National Reading Conference. In Miller, Wallace D. and McNinch, George H. (Eds.) Clemson, South Carolina: The National Reading Conference, Inc., 1976, 29-31.

Tutolo, David. The Study Guide - Types, Purpose and Value. Journal of Reading, 1977, 20, 503-507.

Vacca, Richard. Means of Building Comprehension in Social Studies. In H. L. Herber and P. L. Sanders (Eds.) Research in Reading in the Content Areas: Second Year Report. Syracuse: Syracuse University Press, 1973.

Weaver, W. W. and Kingston, A. J. Questioning in Content Reading. Journal of Reading, 1967, 11, 140-143.

Weintraub, Samuel. What Research Says About Learning to Read. Coordinating Reading Instruction. Glenview, Illinois: Scott, Foresman and Co., 1971.

Research Related to Reflective Reaction Activities

Barbara Flores
Reading Department
University of Arizona

Content area teachers, like most, are interested in facilitating and expanding their students' understanding of what they have read with what they already know. According to Estes and Vaughan (1978), when readers examine relationships between what they know and what they have read, it provides them with an avenue for reflective reaction:

> People remember what they can relate; people comprehend what they can integrate into their cognitive structures. Students need to understand that reading comprehension, if by that we mean understanding what is read, does not stop with the end of the act of reading. It continues as understandings are deepened in reflection (p. 177).

The assumptions and theories about what learning is and how we learn that have predominantly influenced the concept of reading in the content areas can be attributed to Bruner (1965), Ausubel (1968), and Neisser (1969). Preeminently, it is Ausubel's theory of learning that has provided a framework for the investigation of the utility of two specific learning strategies, the advanced organizer and the structured overview. These learning strategies, based on Ausubel's learning theory, have been tested and researched in various instructional settings. This applied research has attempted to test Ausubel's assumptions. This theoretical framework, provided by learning psychologists, has significantly influenced the research and advocacy of certain learning strategies by reading educators, particularly, Herber, Barron, Estes, and Earle (1969). Three basic questions that have

guided the research conducted by learning psychologists and secondary reading educators include: (1) What is learning? (2) How do we learn? and (3) How can we facilitate learning in our teaching?

The purposes of this report are two fold: (1) to restate briefly Ausubel's theoretical framework and assumptions about learning and (2) to present and examine the research on structured overviews that relates to reflective reaction.

Ausubel's Theoretical Framework

It is important that the basic theoretical framework, principles, and assumptions that underlie particular suggested teaching practices be understood. Misinformation, ognorance, or faith in the "experts" is no substitute for basic understanding and knowledge. With this in mind, Ausubel's beliefs about how we learn and the assumptions guiding this theoretical framework need to be synthesized.

Basically Ausubel's theory of learning encompasses two fundamental dimensons of the learning processes: (1) "...the way information is made available to the learner, via either reception or discovery and (2) ...the means, either meaningful or rote, available to the learner to incorporate the new information into his existing cognitive structure which is the organized set of facts, concepts, and generalizations that you have already learned and remember" (Goodwon and Klausmeier, 1975, p. 186).

Another key principle is called subsumption. That is, new information is assimilated into existing cognitive structures. "A new idea is thus, anchored, through subsumption, to a concept already existing for the learner. This anchoring gives the new information stability and provides the mechanism for remembering" (Goodwin and Klausmeier, 1975, p. 186). The underlying assumption that emerges is that the reader's existing knowledge or cognitive

structure is the basis for relating and assimilating new information. This assumption is explicitly stated by Ausubel in the following statement:

> The entire research program has been based on the premise that existing cognitive structure, that is, an individual's organization, stability, and clarity of knowledge in a particular subject matter field at any given time, is the principal factor influencing the learning and retention of meaningful new material. (Ausubel, 1963, p. 217)

Thus, according to Ausubel, the logical rationale for facilitating learning is to "deliberately" try to influence cognitive structuring in order to "maximize" meaningful learning.

As one can see, Ausubel's assumptions and theories have greatly influenced the use and research of the advanced organizer and structured overview, both before reading and after reading. Although, we are primarily addressing how we can facilitate understanding during the reflective reaction phase, one must keep in mind that during reading the reader is interacting with the author and constructing meaning based on his/her past experience, background, and knowledge (Goodman, 1973; Smith, 1971). Bartlett (1932) also found that comprehension or rather understanding was influenced by the reader's prior knowledge, and that this prior knowledge was organized around schemata. Most recently Rumelhart and Ortony (1977) and Anderson, Spiro, and Anderson (1978) have demonstrated that schema theory is an important aspect in the understanding of reading comprehension. Apparently, schemata and cognitive structures are concepts that describe our prior knowledge, experience, and background.

Research Investigations

Research related to the reflective reaction phase is scarce. Very few studies exist that are directly related to studying the affects of post organizers or post structured overviews. For example, the following studies by Barron and Stone (1975), Barron and McCann (1974), Barron and Cooper (1973), Bertou, et.al. (1972), and Graber, et.al. (1972) have researched the efficacy of graphic post organizers in facilitating learning and retention.

Barron and Stone's "...study (1975) addressed the question: what is the effect of student-constructed graphic post organizers upon the learning of vocabulary relationships from a passage in social science?" (p. 173). They hypothesized the (1) the mean scores of the Graphic Post Organizer (GPO) group would be greater than the mean score of the Graphic Advance Organizer (GAO) group and (2) the mean scores of the Graphic Advance Organizer group, as measured by the Mental Health Word Relationships Test, would be greater than the mean score of the control group. This instrument was designed to assess "...the subjects' understanding of hierarchical and parallel relationships which existed among the terminology used in the learning passage" (p. 174). One hundred and forty-one tenth and eleventh graders were randomly assigned to three treatment conditions: (1) the GAO group; (2) the GPO group; and (3) the control group. Each group read a 3000 word passage about mental health, and the topic was assumed to be relatively unfamiliar to most students. This Posttest Only Design did not control for internal threats to validity by determining the students' prior knowledge. This certainly could have influenced the results. The description of the graphic advance organizer appeared to resemble the format of the structured overview. The treatments were administered in a three day period. Both the GAO group and GPO group read the passage twice as did the control group. The GAO group received the graphic advanced

organizer, which the experimenters constructed, prior to reading the passage. The GPO group read the passage first and then in randomly selected groups of two or three constructed their own graphic structured overviews with the vocabulary concepts that were provided. As predicted by the researchers, the GPO group's mean score on the Mental Health Vocabulary Relationships Test was significantly greater at the .05 level. Accepting these results as proof for the significance of using post organizers in the reflective reaction phase must be cautiously considered because we do not know if the test was really testing what it purported to.

Bertou, Clasen, and Lambert (1972) studied the affects of advanced organizers and post organizers and interspersed questions on learning and retention. The one hundred and seventy-six students were selected from a junior high school in Oconomowoc, Wisconsin and were randomly assigned to eight experimental conditions. The topic, which was assumed to be unfamiliar to the students, was about the origins and production of atomic energy. The lessons were video taped. The results showed no significant increase in learning which was evaluated through a multiple choice test whose results were subjected to analysis of variance and to Scheffe post hoc procedures. Accordingly, neither advanced organizers nor post organizers resulted in increased learning. The researchers postulated that this may have been because learning from television is different from learning from written materials.

Graber, Means, and Johnsten (1972) also attempted to determine the affect of advance and post organizers on learning in addition to identifying the students' internal ability to organize learning materials when an external organizer was administered. The sample population was drawn from six sections of an introductory undergraduate chemistry course in a midwestern university. A hundred and forty students participated in a test to determine

organizing ability. The students read a 750 word science oriented passage within a ten minute time limit and then completed an outline structure whereby each had to organize the concepts into their proper sequence in the outline. This activity served as a screening device for selecting students according to their organizing ability. Only those students who scored high and low were used in the experiment. The materials used were the identical ones that Ausubel used in his 1960 study. The results indicated no significant differences among the three types of learning strategies. A 3x2 ANOVA design was used to analyze the results from the learning passage test. "Similary, no significant interaction effect was determined. However, good organizers scored significantly higher than poor organizers" (p. 278). The last finding appears quite predictable in that organizing ability was the criterion for sample selection. The researchers concluded that their findings did not support Ausubel's hypothesis of subsumption. They gave several explanations: (1) difference in the type of students; (2) the difficulty of the subject matter; and/or (3) the lack of sensitivity of the experimental design to pick up the differences.

Discussion, Summary, Conclusions

The studies that related to the use of post reading strategies are very few. Each had a different population, i.e., junior high students, high school students, and college undergraduates. The only study that used the structured overview after reading was Barron and Stone. The lack of control of prior knowledge was not addressed well. Perhaps, it is futile to try to control it since we have varying schemata and cognitive structures. The use of experimental designs to test affects by using measures that did not state their reliability nor validity caused problems; however, this does not mean that we cannot learn about how our students are

learning. Perhaps, we need to examine how students perform in a natural setting as opposed to an experimental one. That is, have students in small groups interacting about what the passage meant and how they related it to what they already knew.

This paper attempted to relate Ausubel's learning theory to the reflective reaction phase and to show the type of research that has been done in this area. As one can tell, the research is limited, but the ideas are intuitively applicable and worth pursuing. More research is certainly needed.

Content area reading teachers can begin to take Ausubel's theoretical principles and begin to test them with their students. They can be researchers along with their students in investigating how they learn, and small group interactions may be an alternative worth exploring. Analyzing how each interrelated what one read with what one knew might serve as an appreciation for the marvel of man's ability to create and expand meaning.

References

Anderson, R. C., Spiro, R. J. and Anderson, M. C. Schemata as scaffolding for the representation of information in connected discourse. American Educational Research Journal, 1978, 15 (3), 433-440.

Ausubel, D. P. The Use of Advance Organizers in the Learning and Retention of Meaningful, Verbal Material. Journal of Educational Psychology, 1960, 51, 267-272.

Ausubel, D. P. Cognitive Structure and the Facilitation of Meaningful Verbal Learning. Journal of Teacher Education, 1963, 14, 217-221.

Ausubel, D. P. The Psychology of Meaningful Verbal Learning. New York: Grune & Stratton, 1963.

Ausubel, D. P. & Fitzgerald, D. The Role of Discriminability in Meaningful Verbal Learning and Retention. Journal of Educational Psychology, 1961, 52, 266-274.

Ausubel, D. P. & Youssef, M. The Role of Discriminability in Meaningful Parallel Learning. Journal of Educational Psychology, 1963, 54, 331-335.

Bertou, D. P., Clasen, R. E. & Lambert, P. An Analysis of the Relative Efficacy of Advanced Organizer, Postorganizers, Interspersed Questions and Combinations Thereof in Facilitating Learning and Retention from a Televised Lecture. Journal of Educational Research, 1972, 65 (7), 329-333.

Barron, R. F. & Stone, V. F. The Effect of Student-Constructed Graphic Post Organizers Upon Learning Vocabulary Relationships. In P. L. Nacke (Ed.) Interaction: Research and Practice in College-Adult Reading. 23rd NRC Yearbook, Clemson, South Carolina: National Reading Conference, 1974.

Barron, R. F. The Use of Vocabulary as an Advance Organizer. In H. L. Herber and P. L. Sanders (Eds.) Research in Reading in the Content Areas: First Year Report. Syracuse: Syracuse University Press, 1969.

Bartlett, F. C. Remembering. Cambridge, England: Cambridge University Press, 1932.

Bruner, J. S. Learning and Thinking. In R. C. Anderson and D. P. Ausubel (Eds.) Readings in the Psychology of Cognition. New York: Holt, Rinehart & Winston, 1965.

Estes, T. H., Mills, D. C., and Barron, R. F. Three Methods of Introducing Students to a Reading-Learning Task in the Content Areas. In H. L. Herber and P. L. Sanders (Eds.) Research in Reading in the Content Areas: First Year Report. Syracuse: Syracuse University Press, 1969.

Estes, T. H. and Vaughan, J. L. Reading and Learning in the Content Classroom: Diagnostic and Instructional Strategies. Boston: Allyn and Bacon Company, 1978.

Earle, R. A. The Use of Vocabulary as a Structured Overview in Seventh Grade Mathematics. In H. L. Herber and P. L. Sanders (Eds.) Research in Reading in the Content Areas: First Year Report. Syracuse: Syracuse University Press, 1969.

Goodman, K. S. Psycholinguistic Universals in the Reading Process. In Frank Smith (Ed.) Psycholinguistics and Reading. New York: Holt, Rinehart and Winston, 1973.

Goodwin, W. L. and Klausmeier, H. J. Facilitating Student Learning: An Introduction to Educational Psychology. New York: Harper & Row, 1975.

Graber, R. A., Means, F. S. and Johnsten, T. D. The Effect of Subsuming Concepts on Student Achievement on Unfamiliar Science Learning Material. Journal of Research in Science Teaching, 9, 277-279, 1972.

Rumelhart, D. E. and Ortony, A. The Representation of Knowledge in Memory. In R. C. Anderson, R. J. Spiro, and W. E. Montague (Eds.) Schooling and the Acquisition of Knowledge. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1977.

Smith, Frank Understanding Reading. New York: Holt, Rinehart and Winston, 1971.

Future Issues

Our next issue is planned for November, 1979, and will focus on the concerns of reading comprehension. Given the recent developments regarding comprehension among linguists and cognitive psychologists, especially those at the Center for the Study of Reading at the University of Illinois in Urbana, the issue of a schematic orientation to reading comprehension has generated substantial interest among reading educators. We intend, therefore, to offer several papers, primarily oriented to theory and research concerns, followed by reactions to those papers that will highlight directions for research with an emphasis on secondary reading. This may be our most industrious effort and, at this point, we believe will be a most provocative publication for secondary reading educators interested in research.

Our fifth issue, scheduled for Spring, 1980, will focus on any research efforts related to secondary reading. Two articles are currently scheduled that provide an historical review of trends in secondary reading programs and a data base for an effective reading in the content areas pre-service and in-service project. This issue is not yet complete and we welcome additional manuscripts.